Cell Research (2010) 20:499-501.

www.nature.com/cr

RESEARCH HIGHLIGHT

Sex matters in the birth of genes

Jessie Colin[1, 2], Domenico Libri[1, 2], Tommaso Villa[1, 2]

[1]LEA Laboratory of Nuclear RNA metabolism, Centre de Génétique Moléculaire, C.N.R.S.-FRE3144, 1, av de la Terrasse, 91190, Gif sur Yvette, France; [2]Centre for mRNP Biogenesis and Metabolism, Department of Molecular Biology, Aarhus University, C.F. Møllers Alle, Bldg. 130, DK-8000 Aarhus C., Denmark
Cell Research (2010) **20**:499-501. doi:10.1038/cr.2010.58; published online 3 May 2010

Recent progress in technology has allowed an extraordinary refinement of our knowledge of transcriptomes, revealing their unexpected complexity. In addition to the large number of siRNAs and miRNAs, many noncoding RNA species are now known to be transcribed by RNA polymerase II (RNAPII), in proximity or overlapping to known transcription units in both sense and antisense orientations, as well as from DNA regions previously thought to be transcriptionally inert or silent. This phenomenon, dubbed 'pervasive' transcription, appears to be evolutionary conserved in all eukaryotes from yeast to human [1]. These noncoding RNAs had previously escaped identification because of their very low abundance in cells grown under standard physiological conditions where they are eliminated by degradative machineries shortly after their transcription by RNAPII. Why cells even produce RNAs destined to degradation is an open question. However, it appears logic that cells use quality control mechanisms to rapidly target potentially hazardous transcripts to degradation while preserving proper expression of coding RNAs. Nevertheless, this is not the whole story, as in budding yeast several examples are beginning to emerge of noncoding RNAs that are involved in gene expression regulation, particularly those arising antisense to ORFs. So far, regulation by noncoding RNAs in *S. cerevisiae* relies on the RNA itself or the very act of its transcription generally leading to inhibition of the target gene through *cis*-acting transcriptional interference mechanisms, or by inducing repressive changes in chromatin structure in interesting analogy to RNAi mechanisms of higher eukaryotes. What these instances of noncoding RNA-mediated regulation tell us is that pervasive transcripts are not just noise but genetic material that can be tested by natural selection and eventually evolve specific functions and be exploited for regulatory purposes.

In April issue of *Cell Research*, Wang and coworkers [2] describe what appears to be a further evolutionary step of a noncoding transcript in budding yeast, that is, acquisition of coding capacity with *de novo* origination of a gene. The authors analyze a sense-antisense regulatory module composed of two protein-coding genes transcribed from the same locus in opposite orientation. Intriguingly, while the antisense *ADF1* (Antisense of Depressing Factor 1) gene shows conservation among hemiascomycete species, the *MDF1* (Mating Depressing Factor 1) gene appears to be transcribed and have coding capability only in *S. cerevisiae*. Rather than the result of gene loss in all the analyzed hemiascomycete species, comparative genomics as well as experimental data favor the hypothesis that the unique presence of *MDF1* in *S. cerevisiae* reflects a new gene resulting from evolution of noncoding sequences. More surprising then is the fact that despite being a "new entry" the Mdf1p protein is shown to be an upstream regulator of the yeast mating pathway. In *S. cerevisiae*, three cell types exist: haploids of either a or α mating type and a/α diploids. Both *MAT* loci produce specific transcriptional regulators governing expression of mating type-specific programs: α1p and α2p in *MATα* cells, a1p and a2p in *MATa* cells. In a/α diploids, α2p represses *MATa*-specific genes but it also physically interacts with a1p to repress transcription of haploid-specific genes. Now, Wang and coworkers show that the *MDF1-ADF1* regulatory module is able to control the mating pathway in haploid α cells (Figure 1). The antisense *ADF1* gene encodes a transcriptional repressor of the sense *MDF1* gene and its inactivation leads to ectopic expression of *MDF1* in haploid α cells. By forming a heterodimer with α2p, Mdf1p leads to repression of haploid-specific genes (including *MATα1*) in α cells, thus blocking the mating pathway and favoring vegetative growth over mating (Figure 1). Because the authors also find that *MDF1* expression confers a selective advantage under vegetative growth

Correspondence: Jessie Colin[a], Tommaso Villa[b]
Tel: +33-1-6982-3830 ext.3799; FAX: +33-1-6982-3877
[a]E-mail: colin@cgm.cnrs-gif.fr
[b]E-mail: villa@cgm.cnrs-gif.fr

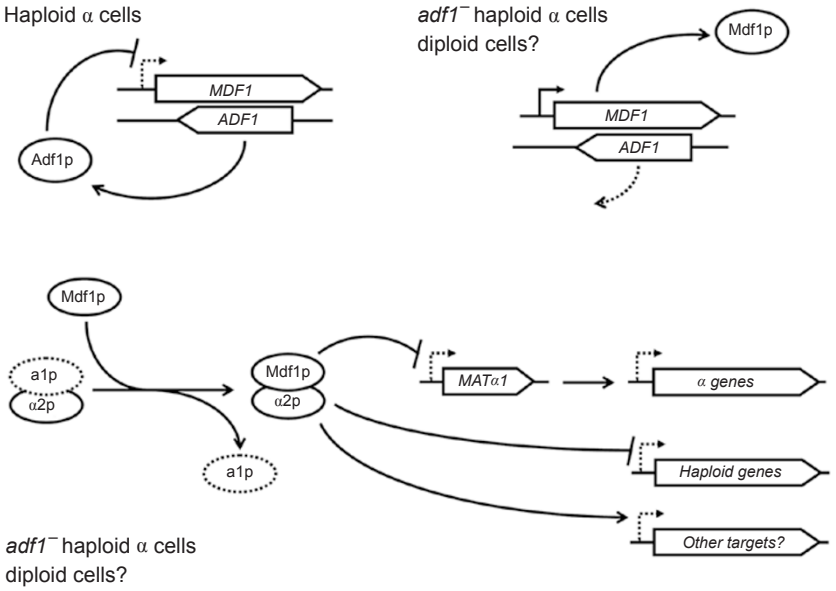

Figure 1 The scheme illustrates the *MDF1-ADF1* sense-antisense regulatory module. Loss of transcriptional repression by Adf1p leads to ectopic expression of *MDF1* in haploid α cells. By forming a heterodimer with α2p, Mdf1p leads to repression of haploid-specific genes (including *MATα1*) in α cells, thus blocking the mating pathway and favoring vegetative growth over mating. It is unknown whether Mdf1p functions in diploids, in which case it is likely that Mdf1p competes with a1p (dotted circles) for α2p binding, possibly regulating expression of additional genes.

conditions, they go on to propose that Mdf1p could enable more flexible shifting between sexual and vegetative growth to best fit the surrounding environmental conditions.

The novelty of the work by Wang and coworkers clearly raises several questions on the mechanisms of *de novo* gene formation as well as on the regulation of the yeast mating pathway. So far, the number of reported *de novo* originated genes with an associated function remains limited. Two *Drosophila* genes, *jingwei* and *sphinx*, originated partly from exogenous coding sequences through retrotransposition events [3, 4]. In contrast, the yeast *MDF1* gene is proposed to have fully originated from noncoding sequences, without any obvious contribution of exogenous DNA. Clearly, the identification of additional examples of noncoding RNAs acquiring coding potential and a specific function will help elucidation of these evolutionary mechanisms. Nevertheless, some yeast noncoding RNAs were found in the cytoplasm associated to polysomes [5], indicating that they might have fortuitously acquired some capacity to associate with components of the translation machinery, maybe as an evolutionary step towards a coding function.

Intriguingly, transitioning from vegetative to sexual growth in budding yeast is also regulated by a different sense-antisense module, that formed by the *IME4* gene and, in this case, its antisense noncoding RNA [6]. *IME4* is required for entry in meiosis in a/α diploids. In haploid cells, constitutive transcription of the antisense noncoding RNA prevents expression of the *IME4* gene by a transcriptional interference mechanism. In diploid cells, the a1p/α2p heterodimer represses transcription of the antisense noncoding RNA, thus permitting transcription of *IME4* and eventually the switch from vegetative growth to meiosis. It is fascinating that two major events in the life cycle of *S. cerevisiae*, the mating pathway and entry into meiosis, have co-opted sense-antisense regulatory modules and exploited them for controlling key steps of these pathways. On the other hand, just because of the importance of these events for yeast fitness, it is reasonable that multiple overlapping strategies are at work to fine-tune regulation. Indeed, recent data nicely illustrated the existence of a trade-off between growth rate and mating efficiency, where reductions in the latter result in a growth rate increase and thus an advantageous adaptation to the environment [7]. Consistently, we might expect *S. cerevisiae* to accommodate possibly more examples of these regulatory circuitries in the context of its sexual cycle. Curiously enough, it is of note that one of the *de novo* originated fly genes, *sphinx*, is also involved in sexual reproduction control, specifically in *Drosophila* male courtship, maybe indicating an evolutionary conserved need to modulate sexual pathways.

Of course, several questions regarding the *MDF1-ADF1* regulatory module remain open. Notably, because of its evolutionary conservation, it will be important to determine whether Adf1p shares common targets in the different hemiascomycete species or if its function has co-evolved with the appearance of *MDF1* in *S. cerevisiae*. In this regard, it will be interesting to determine whether Adf1p can repress transcription of ancestor noncoding sequences of *MDF1* in other hemiascomycetes, maybe explaining the lack of *MDF1* RNA detection in other yeasts. Furthermore, as Mdf1p heterodimerizes with α2p, it is fully possible that Mdf1p competes with a1p for α2p binding in diploids, thus impacting the transcriptional output (Figure 1). In this regard, it will be worthwhile to analyze the possible impact of this Mdf1p/α2p regulator on the *IME4* sense-antisense

circuit. In any case, future studies will tell us more about the possible evolution of noncoding RNAs towards novel coding genes.

References

1 Jacquier A. The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat Rev Genet* 2009; **10**:833-844.

2 Li D, Dong Y, Jiang Y, *et al.* A *de novo* originated gene depresses budding yeast mating pathway and is repressed by the protein encoded by its antisense strand. *Cell Res* 2010; **20**:408-420.

3 Zhang J, Dean AM, Brunet F, Long M. Evolving protein functional diversity in new genes of *Drosophila*. *Proc Natl Acad Sci USA* 2004; **101**:16246-16250.

4 Dai H, Chen Y, Chen S, *et al.* The evolution of courtship behaviors through the origination of a new gene in *Drosophila*. *Proc Natl Acad Sci USA* 2008; **105**:7478-7483.

5 Thompson DM, Parker R. Cytoplasmic decay of intergenic transcripts in *Saccharomyces cerevisiae*. *Mol Cell Biol* 2007; **27**:92-101.

6 Hongay CF, Grisafi PL, Galitski T, Fink GR. Antisense transcription controls cell fate in *Saccharomyces cerevisiae*. *Cell* 2006; **127**:735-745.

7 Lang GI, Murray AW, Botstein D. The cost of gene expression underlies a fitness trade-off in yeast. *Proc Natl Acad Sci USA* 2009; **106**:5755-5760.